A Content-Aware Image Prior

Taeg Sang Cho[†], Neel Joshi[‡], C. Lawrence Zitnick[‡], Sing Bing Kang[‡], Richard Szeliski[‡], William T. Freeman[†]
[†] Massachusetts Institute of Technology, [‡] Microsoft Research

Abstract

In image restoration tasks, a heavy-tailed gradient distribution of natural images has been extensively exploited as an image prior. Most image restoration algorithms impose a sparse gradient prior on the whole image, reconstructing an image with piecewise smooth characteristics. While the sparse gradient prior removes ringing and noise artifacts, it also tends to remove mid-frequency textures, degrading the visual quality. We can attribute such degradations to imposing an incorrect image prior. The gradient profile in fractal-like textures, such as trees, is close to a Gaussian distribution, and small gradients from such regions are severely penalized by the sparse gradient prior.

To address this issue, we introduce an image restoration algorithm that adapts the image prior to the underlying texture. We adapt the prior to both low-level local structures as well as mid-level textural characteristics. Improvements in visual quality is demonstrated on deconvolution and denoising tasks.

1. Introduction

Image enhancement algorithms resort to image priors to hallucinate information lost during the image capture. In recent years, image priors based on image gradient statistics have received much attention. Natural images often consist of smooth regions with abrupt edges, leading to a heavy-tailed gradient profile. We can parameterize heavy-tailed gradient statistics with a generalized Gaussian distribution or a mixture of Gaussians. Prior works hand-select parameters for the model distribution, and fix them for the entire image, imposing the same image prior everywhere [10, 17, 23]. Unfortunately, different textures have different gradient statistics even within a single image, therefore imposing a single image prior for the entire image is inappropriate (Figure 1).

We introduce an algorithm that adapts the image prior to both low-level local structures as well as mid-level texture cues, thereby imposing the correct prior for each texture. [1] Adapting the image prior to the image content improves the quality of restored images.

[1] Strictly speaking, an estimate of image statistics made after examining the image is no longer a "prior" probability. But the fitted gradient distributions play the same role as an image prior in image reconstruction equations, and we keep that terminology.

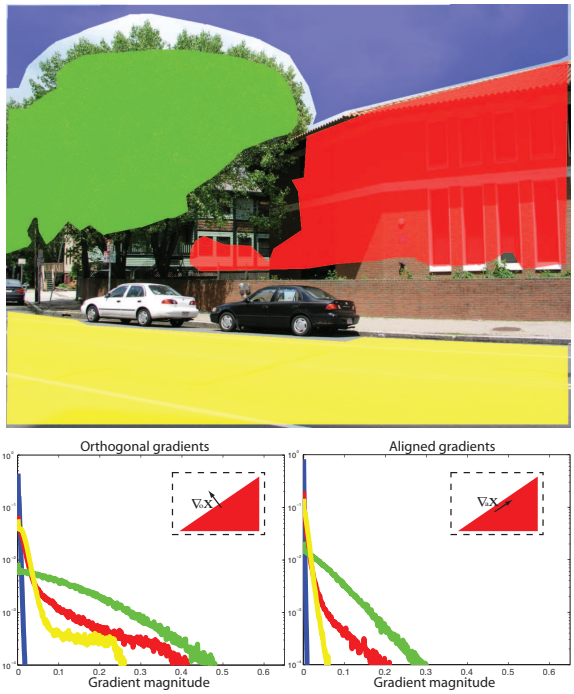

Figure 1. The colored gradient profiles correspond to the region with the same color mask. The steered gradient profile is spatially variant in most natural images. Therefore, the image prior should adapt to the image content. Insets illustrate how steered gradients adapt to local structures.

2. Related work

Image prior research revolves around finding a good image transform or basis functions such that the transformed image exhibits characteristics distinct from unnatural images. Transforms derived from signal processing have been exploited in the past, including the Fourier transform [12], the wavelet transform [28], the curvelet transform [6], and the contourlet transform [8].

Basis functions learned from natural images have also been introduced. Most techniques learn filters that lie in the null-space of the natural image manifold [20, 30, 31, 32]. Aharon *et al*. [1] learns a vocabulary from which a natural image is composed. However, none of these techniques adapt the basis functions to the image under analysis.

Edge-preserving smoothing operators do adapt to local structures while reconstructing images. The anisotropic diffusion operator [5] detects edges, and smoothes along edges

but not across them. A similar idea appeared in a probabilistic framework called a Gaussian conditional random field [26]. A bilateral filter [27] is also closely related to anisotropic operators. Elad [9] and Barash [2] discuss relationships between edge-preserving operators.

Some image models adapt to edge orientations as well as magnitudes. Hammond *et al*. [13] present a Gaussian scale mixture model that captures the statistics of gradients adaptively steered in the dominant orientation in image patches. Roth *et al*. [21] present a random field model that adapts to the oriented structures. Bennett *et al*. [3] and Joshi *et al*. [14] exploit a prior on colors: an observation that there are two dominant colors in a small window.

Adapting the image prior to textural characteristics was investigated for gray-scale images consisting of a single texture [23]. Bishop *et al*. [4] present a variational image restoration framework that breaks an image into square blocks and adapts the image prior to each block independently (i.e. the image prior is fixed within the block). However, Bishop *et al*. [4] do not address issues with estimating the image prior at texture boundaries.

3. Image characteristics

We analyze the statistics of gradients adaptively steered in the dominant local orientation of an image x. Roth *et al*. [21] observe that the gradient profile of orthogonal gradients $\nabla_o x$ is typically of higher variance compared to that of aligned gradients $\nabla_a x$, and propose imposing different priors on $\nabla_o x$ and $\nabla_a x$. We show that different textures within the same image also have distinct gradient profiles.

We parameterize the gradient profile using a generalized Gaussian distribution:

$$p(\nabla x) = \frac{\gamma \lambda^{(\frac{1}{\gamma})}}{2\Gamma(\frac{1}{\gamma})} \exp(-\lambda \|\nabla x\|^{\gamma}) \qquad (1)$$

where Γ is a Gamma function, and γ, λ are the shape parameters. γ determines the peakiness and λ determines the width of a distribution. We assume that $\nabla_o x$ and $\nabla_a x$ are independent: $p(\nabla_o x, \nabla_a x) = p(\nabla_o x)p(\nabla_a x)$.

3.1. Spatially variant gradient statistics

The local gradient statistics can be different from the global gradient statistics. Figure 1 shows the gradient statistics of the colored regions. Two phenomena are responsible for the spatially variant gradient statistics: the material and the viewing distance. For example, a building is noticeably more piecewise smooth than a gravel path due to material properties, whereas the same gravel path can exhibit different gradient statistics depending on the viewing distance. To account for the spatially variant gradient statistics, we propose adjusting the image prior locally.

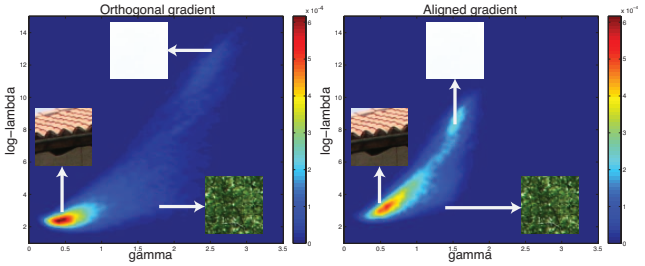

Figure 2. The distribution of $\gamma, \ln(\lambda)$ for $\nabla_o x, \nabla_a x$ in natural images. While the density is the highest around $\gamma = 0.5$, the density tails off slowly with significant density around $\gamma = 2$. We show, as insets, some patches from Figure 1 that are representative of different $\gamma, \ln(\lambda)$.

3.2. The distribution of γ, λ in natural images

Different textures give rise to different gradient profiles and thus different γ, λ. We study the distribution of the shape parameters γ, λ in natural images. We sample $\sim 110,000$ patches of size 41×41 from 500 high quality natural images. We fit the gradient profile from each patch to a generalized Gaussian distribution to associate each patch with γ, λ. We fit the distribution by searching for γ, λ that minimize the Kullback-Leibler (KL) divergence between the empirical gradient distribution and the model distribution p, which is equivalent to minimizing the negative log-likelihood of the model distribution evaluated over the gradient samples:

$$[\tilde{\gamma}, \tilde{\lambda}] = \underset{\gamma,\lambda}{\operatorname{argmin}} \left\{ -\frac{1}{N} \sum_{i=1}^{N} \ln\left(p(\nabla x_i)\right) \right\} \qquad (2)$$

Figure 2 shows the Parzen-window fit to sampled $\tilde{\gamma}, \ln(\tilde{\lambda})$ for $\nabla_o x, \nabla_a x$. For orthogonal gradients $\nabla_o x$, there exists a large cluster near $\gamma = 0.5, \ln(\lambda) = 2$. This cluster corresponds to patches from a smooth region with abrupt edges or a region near texture boundaries. This observation supports the dead leaves image model – an image is a collage of overlapping instances [16, 19]. However, we also observe a significant density even when γ is greater than 1. Samples near $\gamma = 2$ with large λ correspond to flat regions such as sky, and samples near $\gamma = 2$ with small λ correspond to fractal-like textures such as tree leaves or grass. We observe similar characteristics for aligned gradients $\nabla_a x$ as well. The distribution of shape parameters suggests that a significant portion of natural images is not piecewise smooth, which justifies adapting the image prior to the image content.

4. Adapting the image prior

The goal of this paper is to identify the correct image prior (i.e. shape parameters of a generalized Gaussian distribution) for each pixel in the image. One way to identify the image prior is to fit gradients from every sliding window to a generalized Gaussian distribution (Eq 2), but the

required computation would be overwhelming. We introduce a regression-based method to estimate the image prior.

4.1. Image model

Let y be an observed degraded image, k be a blur kernel (a point-spread function or a PSF), and x be a latent image. Image degradation is modeled as a convolution process:

$$y = k \otimes x + n \tag{3}$$

where $\otimes$ is a convolution operator, and n is an observation noise. The goal of a (non-blind) image restoration problem is to recover a clean image x from a degraded observation y given a blur kernel k and a standard deviation of noise η, both of which can be estimated through other techniques [10, 18].

We introduce a conditional random field (CRF) model to incorporate texture variations within the image restoration framework. Typically, a CRF restoration model can be expressed as follows:

$$p(x|y,k,\eta) = \frac{1}{M}\prod_i \phi_y(x;y_i,k,\eta)\phi_x(x) \tag{4}$$

where M is a partition function and i is a pixel index. ϕ_y is derived from the observation process Eq 3; ϕ_x from the assumed image prior:

$$\phi_y(x;y_i,k,\eta) \propto \exp(-\frac{(y_i-(k\otimes x)_i)^2}{2\eta^2}) \tag{5}$$

$$\phi_x(x) \propto \exp(-\lambda\|\nabla x\|^\gamma) \tag{6}$$

To model the spatially variant gradient statistics, we introduce an additional hidden variable z, called *texture*, to the conventional CRF model. z controls the shape parameters of the image prior:

$$p(x,z|y,k,\eta) = \frac{1}{M}\prod_i \phi_y(x;y_i,k,\eta)\phi_{x,z}(x,z) \tag{7}$$

where $\phi_{x,z}(x,z) \propto \exp(-\lambda(z)\|\nabla x\|^{\gamma(z)})$. We model z as a continuous variable since the distribution of $[\gamma,\lambda]$ is heavy-tailed and does not form tight clusters (Figure 2).

We maximize $p(x|y,k,\eta)$ to estimate a clean image $\hat{x}$. To do so, we approximate $p(x|y,k,\eta)$ by the function $p(x,z|y,k,\eta)$ at the mode $\hat{z}$:

$$p(x|y,k,\eta) = \int_z p(x,z|y,k,\eta)dz \approx p(x,\hat{z}|y,k,\eta) \tag{8}$$

Section 4.2 discusses how we estimate $\hat{z}$ for each pixel.

4.2. Estimating the texture $\hat{z}$

A notable characteristic of a zero-mean generalized Gaussian distribution is that the variance v and the fourth moment f completely determine the shape parameters $[\gamma,\lambda]$ [24]:

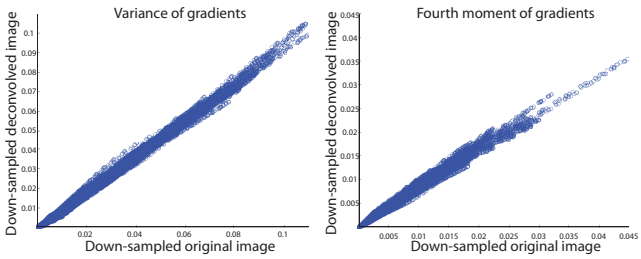

Figure 3. The local variance and fourth moments of gradients computed from the deconvolved, down-sampled image of Figure 1 are closely correlated with those of the down-sampled *original* image.

$$v = \frac{\Gamma(3/\gamma)}{\lambda^{\frac{2}{\gamma}}\Gamma(1/\gamma)}, f = \frac{\Gamma(5/\gamma)}{\lambda^{\frac{4}{\gamma}}\Gamma(1/\gamma)} \tag{9}$$

To take advantage of these relationships, we define the local texture around a pixel i, $\hat{z}_i$, as a two dimensional vector. The first dimension is the variance v_i of gradients in the neighborhood of a pixel i, and the second dimension is the fourth moment f_i of gradients in the neighborhood of a pixel i:

$$\hat{z}_i = [v_i(\nabla x), f_i(\nabla x)] \tag{10}$$

Qualitatively, the variance of gradients $v_i(\nabla x)$ encodes the width of the distribution, and the fourth moment $f_i(\nabla x)$ encodes the peakiness of the distribution. Note that we can easily compute v_i, f_i by convolving the gradient image with a window that defines the neighborhood. We use a Gaussian window with a 4-pixel standard deviation.

Estimating the texture $\hat{z}$ from the observation y The texture $\hat{z}$ should be estimated from the sharp image x we wish to reconstruct, but x is not available when estimating $\hat{z}$. We address this issue by estimating the texture $\hat{z}$ from an image reconstructed using a spatially invariant image prior. We hand-select the spatially invariant prior with a weak gradient penalty so that textures are reasonably restored: $[\gamma_o = 0.8, \lambda_o = 6.5], [\gamma_a = 0.6, \lambda_a = 6.5]$. A caveat is that the fixed prior deconvolution may contaminate the gradient profile of the reconstructed image, which may induce texture estimation error. To reduce such deconvolution noise, we down-sample the deconvolved image by a factor of 2 in both dimensions before estimating the texture $\hat{z}$. The gradient profile of natural images is often scale invariant due to fractal properties of textures and piecewise smooth properties of surfaces [16, 19], whereas that of the deconvolution noise tends to be scale variant. Therefore, the texture $\hat{z}$ estimated from the down-sampled deconvolved image better resembles the texture of the original sharp image.

4.3. Estimating the shape parameters γ, λ from $\hat{z}$

We could numerically invert Eq 9 to directly compute the shape parameters $[\gamma,\lambda]$ from the variance and fourth moment estimates [24]. However, a numerical inversion is computationally expensive and is sensitive to noise. We instead

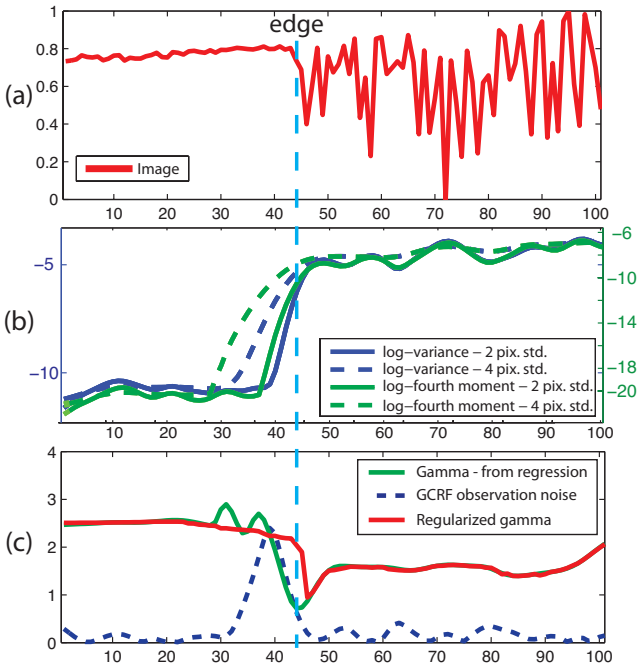

Figure 4. We regularize the estimated shape parameters using a GCRF such that the texture transition mostly occurs at the texture boundary. We model the observation noise in the GCRF as the *variance* of the variance and fourth moments estimated from two Gaussian windows with different standard deviations – 2-pixel and 4-pixel, as shown in (b). This reduces the shape parameter estimation error at texture boundaries, as shown in (c) (compare green and red curves).

use a kernel regressor that maps the log of the texture $\ln(\hat{z})$ to shape parameters $[\gamma, \ln(\lambda)]$.

The regressor should learn the mapping from the texture $\hat{z}$ of the down-sampled *deconvolved* image to shape parameters in order to account for any residual deconvolution noise in the estimated texture $\hat{z}$. Since the deconvolved image, thus $\hat{z}$, depends on the blur kernel and the noise level, we would have to train regressors discriminatively for each degradation scenario, which is intractable. However, we empirically observe in Figure 3 that the variance and fourth moment of the deconvolved, down-sampled image are similar to those of the down-sampled original image. Therefore, we learn a *single* regressor that maps the variance and fourth moment of the down-sampled *original* image to the shape parameters, and use it to estimate the shape parameters from the down-sampled deconvolved image.

To learn the regression function, we sample $125,000$ patches of size 17×17 pixels from 500 high quality natural images. We fit the gradient profile of each patch to a generalized Gaussian distribution, and associate each fit with the variance and fourth moment of gradients in the down-sampled version of each patch (9×9 pixels). We use the collected data to learn the mapping from $[\ln(v), \ln(f)]$ to $[\gamma, \ln(\lambda)]$ using LibSVM [7]. We use a 10-fold cross validation technique to avoid over-fitting.

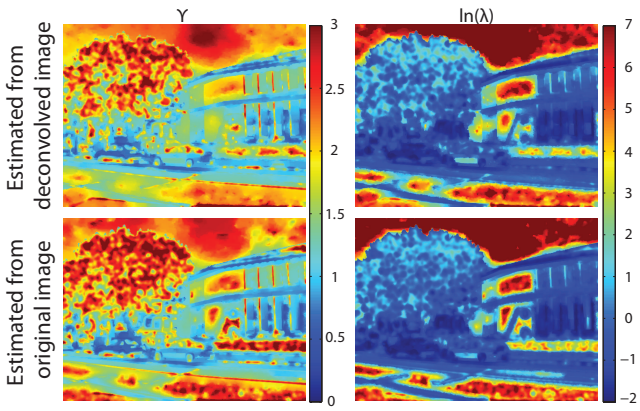

Figure 5. The shape parameters for orthogonal gradients, estimated from the down-sampled deconvolved image of Figure 1 (reconstructed from an image degraded with 5% noise and the blur in Figure 7) and the down-sampled original image. The shape parameters estimated from the deconvolved image are similar to those of the original image.

4.4. Handling texture boundaries

If multiple textures appear within a single window, the estimated shape prior can be inaccurate. Suppose we want to estimate the image prior for a 1-dimensional slice of an image (Figure 4(a)). Ideally, we should recover two regions with distinct shape parameters that abut each other via a thin band of shape parameters corresponding to an edge. However, the estimated image prior becomes "sparse" (i.e. small γ) near the texture boundary even if pixels do not correspond to an edge (the green curve in Figure 4(c)). The use of a finite-size window for computing v and f causes this issue.

To recover shape parameters near texture boundaries, we regularize the estimated shape parameters using a Gaussian conditional random field (GCRF) [26]. Essentially, we want to smooth shape parameters only near texture boundaries. A notable characteristic at texture boundaries is that $\hat{z}$'s estimated from two different window sizes tend to be different from each other: while a small window spans a homogenous texture, a larger window could span two different textures, generating different $\hat{z}$'s. We leverage this characteristic to smooth only near texture boundaries by defining the observation noise level in GCRF as a *mean variance* of the variance v and of the fourth moment f estimated from windows of two different sizes (Gaussian windows with 2-pixel and 4-pixel standard deviations.) If the variance of these estimates is large, the central pixel is likely to be near a texture boundary, thus we smooth the shape parameter at the central pixel. Appendix A discusses the GCRF model in detail. Figure 4(c) shows the estimated shape parameters before and after regularization along with the estimated GCRF observation noise. After regularization, two textures are separated by a small band of sparse image prior corresponding to an edge.

Figure 5 shows the estimated shape parameters for or-

thogonal gradients of Figure 1. In the top row of Figure 5, the parameters are estimated from the image reconstructed from 5% noise and the blur in Figure 7. We observe that the estimated prior in the tree region is close to Gaussian (i.e. $\gamma = 2 \sim 3$), whereas the estimated prior in the building region is sparse (i.e. $\gamma < 1$). The estimated shape parameters are similar to parameters estimated from the down-sampled, original image (the bottom row of Figure 5). This supports the claim that shape parameters estimated from a degraded input image reasonably accurate.

4.5. Implementation details

We minimize the negative log-posterior to reconstruct a clean image $\hat{x}$:

$$\hat{x} = \underset{x}{\operatorname{argmin}}\{\frac{(y - k \otimes x)^2}{2\eta^2} + w \sum_{i=1}^{N} (\lambda_o(\hat{z}_i)\|\nabla_o x(i)\|^{\gamma_o(\hat{z}_i)} + \lambda_a(\hat{z}_i)\|\nabla x\|^{\gamma_a(\hat{z}_i)})\} \quad (11)$$

where $[\gamma_o, \lambda_o], [\gamma_a, \lambda_a]$ are estimated parameters for orthogonal and aligned gradients, respectively, and w is a weighting term that controls the gradient penalty. $w = 0.025$ in all examples. We minimize Eq 11 using an iterative reweighted least squares algorithm [17, 25].

5. Experimental results

We evaluate the performance of the content-aware image prior for deblurring and denoising tasks. We compare our results to those reconstructed using a sparse unsteered gradient prior [17] and a sparse steered gradient prior [21], using peak signal-to-noise ratio (PSNR) and gray-scale structural similarity (SSIM) [29] as quality metrics. We augmented the steerable random fields [21], which introduced denoising and image inpainting as applications, to perform deconvolution using the sparse steered gradient prior. In all experiments, we use the first order and the second order gradient filters [11]. We can augment these algorithms using higher order gradient filters to improve reconstruction qualities, but it is not considered in this work. The test set consists of 21 high quality images downloaded from LabelMe [22] with enough texture variations within each image.

Non-blind deconvolution The goal of non-blind deconvolution is to reconstruct a sharp image from a blurred, noisy image given a blur kernel and a noise level. We generate our test set by blurring images with the kernel shown in Figure 7, and adding 5% noise to blurred images. Figure 6 shows the measured PSNR and SSIM for different deconvolution methods. The content-aware prior deconvolution method performs favorably compared to the competing methods, both in terms of PSNR and SSIM. The benefit of using a spatially variant prior is more pronounced for images with large textured regions. If the image consists primarily

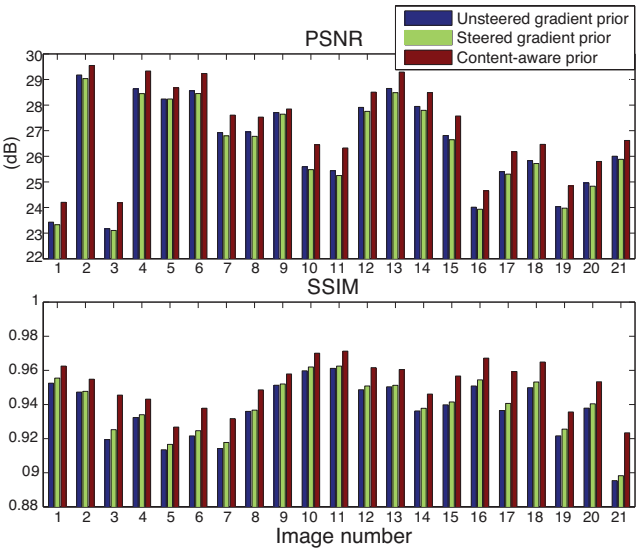

Figure 6. Image deconvolution results : PSNR and SSIM. Mean PSNR: unsteered gradient prior – 26.45 dB, steered gradient prior – 26.33 dB, **content-aware prior – 27.11 dB**. Mean SSIM: unsteered gradient prior – 0.937, steered gradient prior – 0.940, **content-aware prior – 0.951**.

of piecewise smooth objects such as buildings, the difference between the content-aware image prior and others is minor. Figure 7 compares the visual quality of images reconstructed using different priors. We observe that textured regions are best reconstructed using the content-aware image prior, illustrating the benefit of adapting the image prior to textures.

Denoising The goal of denoising is to reconstruct a sharp image from a noisy observation given a noise level. We consider reconstructing clean images from degraded images at two noise levels: 5% and 10%. Figure 8 shows the measured PSNR and SSIM for the denoising task. When the noise level is low (5%), the content-aware prior reconstructs images with lower PSNR compared to competing methods. One explanation is that the content-aware prior may not remove all the noise in textured regions (such as trees) because the gradient statistics of noise is similar to that of the underlying texture. Such noise, however, does not disturb the visual quality of textured regions. The SSIM measure, which is better correlated with the perceptual quality [29], shows that the content-aware image prior performs slightly worse, if not comparably, compared to other methods at a 5% noise level. It's worth noting that when the noise level is low, the observation term is strong so that reconstructed images do not depend heavily on the image prior. The top row of Figure 9 shows that at a 5% noise level, reconstructed images are visually similar.

When the noise level is high (10%), SSIM clearly favors images reconstructed using the content-aware prior. In this case, the observation term is weak, thus the image prior plays an important role in the quality of reconstructed im-

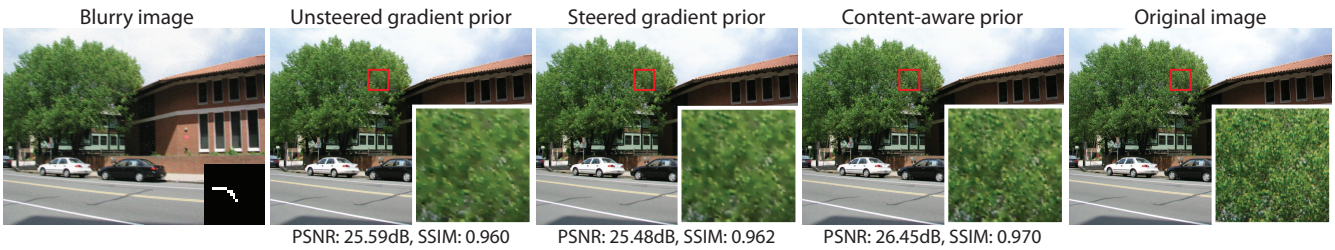

Figure 7. Adapting the image prior to textures leads to better reconstructions. The red box denotes the cropped region.

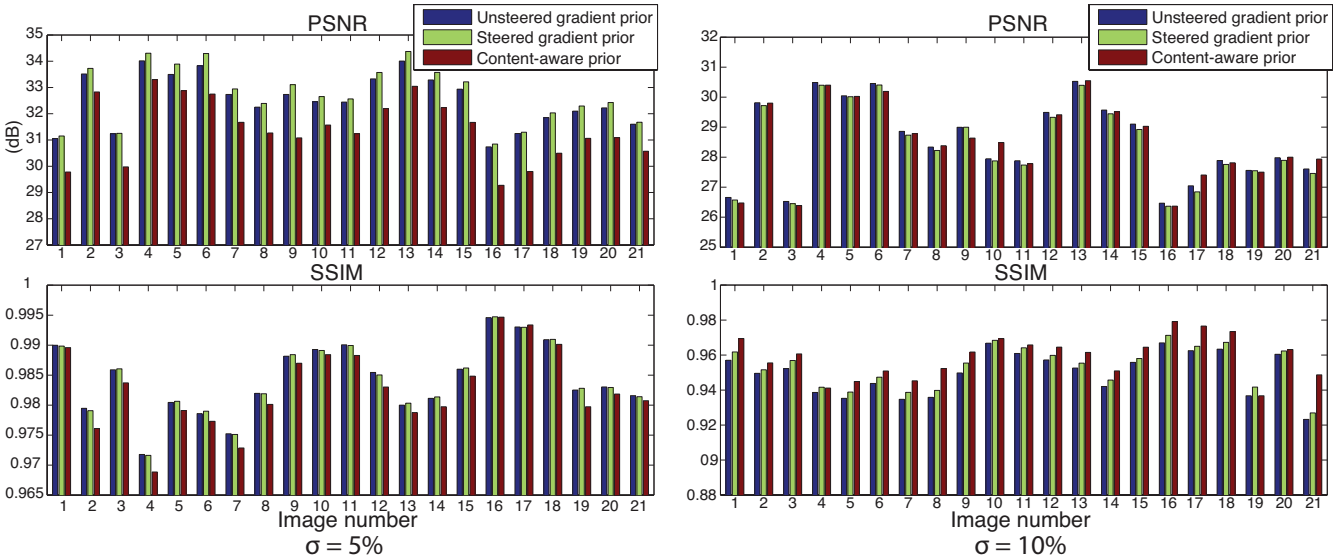

Figure 8. Image denoising results : PSNR and SSIM. At 5% noise = Mean PSNR: unsteered gradient prior – 32.53 dB, steered gradient prior – 32.74 dB, **content-aware prior – 31.42 dB**. Mean SSIM: unsteered gradient prior – 0.984, steered gradient prior – 0.984, **content-aware prior – 0.982**. At 10% noise = Mean PSNR: unsteered gradient prior – 28.54 dB, steered gradient prior – 28.43 dB, **content-aware prior – 28.52 dB**. Mean SSIM: unsteered gradient prior – 0.950, steered gradient prior – 0.953, **content-aware prior – 0.959**

ages. Therefore, the performance benefit from using the content-aware prior is more pronounced. The bottom row of Figure 9 shows denoising results at a 10% noise level, supporting our claim that the content-aware image prior generates more visually pleasing textures.

Figure 10 shows the result of deconvolving a real blurry image captured with a handheld camera. We estimate the blur kernel using the algorithm in Fergus *et al*. [10]. Again, textured regions are better reconstructed using our method.

5.1. User study

We conducted a user study on Amazon Mechanical Turk to compare the visual quality of the reconstructed images. We evaluated 5 randomly selected images for each degradation scenario considered above. Each user views two images, one reconstructed using the content-aware prior and another reconstructed using either the unsteered gradient prior or the steered gradient prior. The user selects the visually pleasing one of the two, or selects *"There is no difference"* option. We cropped each test image to 500×350 pixels to ensure that the users view both images without scrolling. We did not refer to restored images while selecting crop regions in order to minimize bias.

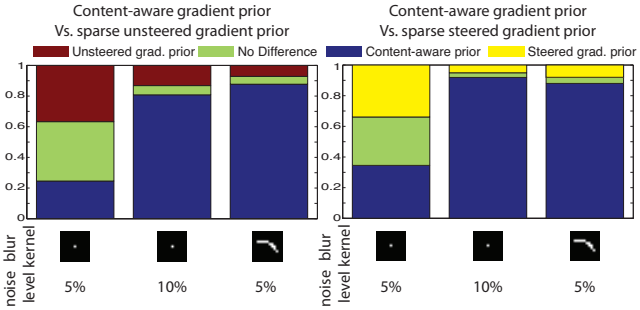

Figure 11. User study results. The blue region corresponds to the fraction of users that favored our reconstructions for each degradation scenario. At a low degradation level, users do not prefer one method over another, but as the level of degradation increases, users clearly favor the content-aware image prior.

We gathered about 20 user opinions for each comparison. In Figure 11, we show the average user preference in each degradation scenario. Users did not have a particular preference when the degradation was small (e.g. 5% noise), but at a high image degradation level users clearly favored the content-aware image prior over others.

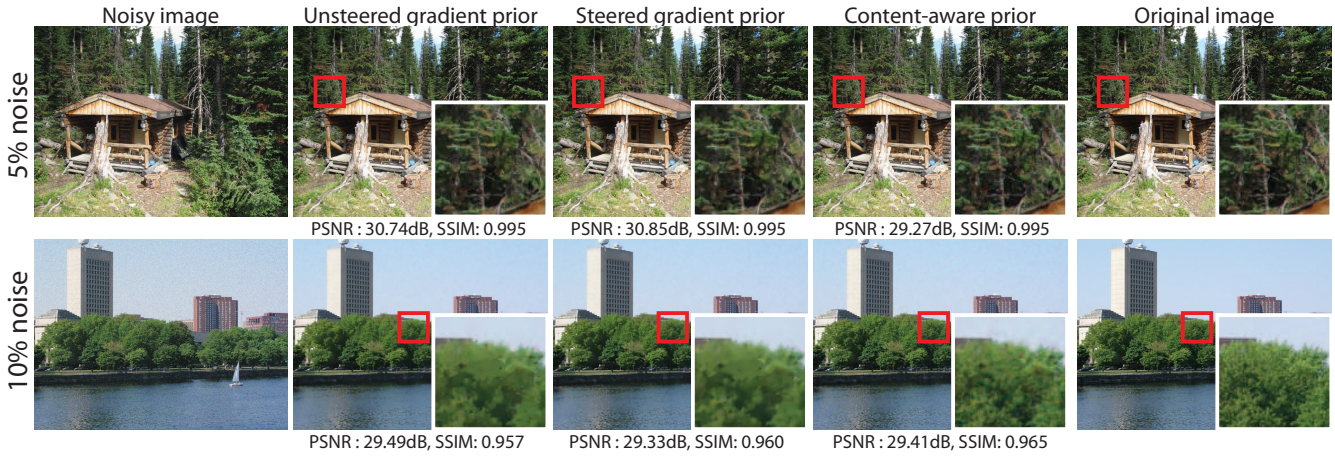

Figure 9. The visual comparisons of denoised images. The red box denotes the cropped region. At 5% noise level, while the PSNR of our result is lower than those of competing algorithms, visually the difference is imperceptible. At 10% noise level, content-aware prior outperforms others in terms of both the PSNR and the SSIM, and is more visually pleasing.

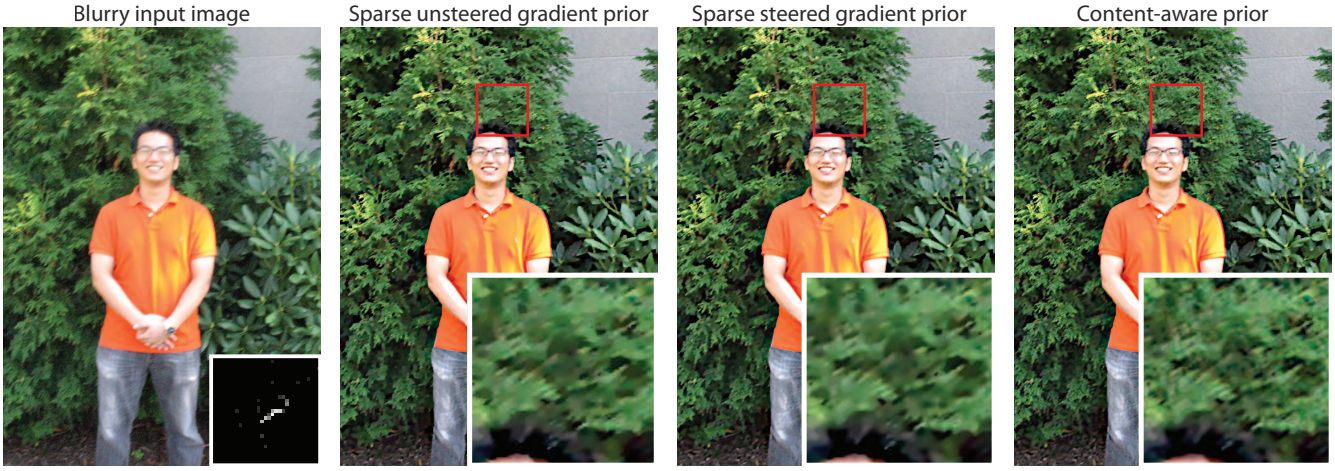

Figure 10. The deconvolution of a blurred image taken with a hand-held camera. We estimate the blur kernel using Fergus *et al*. [10]. The red box denotes the cropped region. The textured region is better reconstructed using the content-aware image prior.

5.2. Discussions

A limitation of our algorithm, which is shared with algorithms using a conditional random field model with hidden variables [14, 21, 26], is that hidden variables, such as the magnitude and/or orientation of an edge, or textureness of a region, are estimated from the degraded input image or the image restored through other means. Any error from this preprocessing step induces error in the final result.

Another way to estimate a spatially variant prior is to segment the image into regions and assume a single prior within each segment. Unless we segment the image into many pieces, the estimated prior can be inaccurate. Also, the segmentation may inadvertently generate artificial boundaries in reconstructed images. Therefore, we estimate a distinct image prior for each pixel in the image.

6. Conclusion

We have explored the problem of estimating spatially variant gradient statistics in natural images, and exploited the estimated gradient statistics to adaptively restore different textural characteristics in image restoration tasks. We show that the content-aware image prior can restore piecewise smooth regions without over-smoothing textured regions, improving the visual quality of reconstructed images as verified through user studies. Adapting to textural characteristics is especially important when the image degradation is significant.

Appendix A

Gaussian CRF model for $[\gamma, \lambda]$ regularization

We regularize the regressor outputs $[\tilde{\gamma}, \tilde{\lambda}]$ using a Gaussian Conditional Random Fields (GCRF). We maximize the following probability to estimate regularized γ:

$$P(\gamma; \tilde{\gamma}) \propto \prod_{i,j \in \mathbb{N}(i)} \psi(\tilde{\gamma}_i | \gamma_i) \Psi(\gamma_i, \gamma_j) \tag{12}$$

where $\mathbb{N}(i)$ denotes the neighborhood of i, ψ is the observation model and Ψ is the neighborhood potential:

$$\psi(\tilde{\gamma_i}|\gamma_i) \propto \exp\left(-\frac{(\tilde{\gamma_i}-\gamma_i)^2}{2\sigma_l^2}\right)$$
$$\Psi(\gamma_i,\gamma_j) \propto \exp\left(-\frac{(\gamma_i-\gamma_j)^2}{2\sigma_n^2(i,j)}\right) \quad (13)$$

We set σ_l and σ_n adaptively. We set the variance $\sigma_n^2(i,j)$ of the neighboring γ estimates γ_i, γ_j as $\sigma_n^2(i,j) = \alpha(x(i) - x(j))^2$, where x is the image and $\alpha = 0.01$ controls how smooth the neighboring estimates should be. σ_n encourages the discontinuity at strong edges of the image x [26]. The observation noise σ_l^2 is the average variance of the variance and fourth moment estimates (for two Gaussian windows with standard deviation = 2 pixels, 4 pixels). We use the same GCRF model to regularize $\ln(\lambda)$ with $\alpha = 0.001$.

Acknowledgment

This work was done while the first author was an intern at MSR. Funding for this work was provided by NGA NEGI-1582-04-0004, MURI Grant N00014-06-1-0734, and gifts from Microsoft, Google and Adobe. The first author is also supported by Samsung Scholarship Foundation.

References

[1] M. Aharon, M. Elad, and A. Bruckstein. The K-SVD: an algorithm for designing of overcomplete dictionaries for sparse representation. *IEEE TSP*, Nov. 2006. 1

[2] D. Barash. A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation. *IEEE TPAMI*, 24(6):844 – 847, June 2002. 2

[3] E. P. Bennett, M. Uyttendaele, C. L. Zitnick, R. Szeliski, and S. B. Kang. Video and image bayesian demosaicing with a two color image prior. In *ECCV*, 2006. 2

[4] T. E. Bishop, R. Molina, and J. R. Hopgood. Nonstationary blind image restoration using variational methods. In *IEEE ICIP*, 2007. 2

[5] M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger. Robust anisotropic diffusion. *IEEE TIP*, Mar. 1998. 1

[6] E. J. Candes and D. L. Donoho. Curvelets - a surprisingly effective nonadaptive representation for objects with edges, 1999. 1

[7] C.-C. Chang and C.-J. Lin. *LIBSVM: a library for support vector machines*, 2001. 4

[8] M. N. Do and M. Vetterli. The contourlet transform: an efficient directional multiresolution image representation. *IEEE TIP*, Dec. 2005. 1

[9] M. Elad. On the origin of the bilateral filter and ways to improve it. *IEEE TIP*, Oct. 2002. 2

[10] R. Fergus, B. Singh, A. Hertzmann, S. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM TOG (SIGGRAPH)*, 2006. 1, 3, 6, 7

[11] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE TPAMI*, 13(9), Sept. 1991. 5

[12] Gonzalez and Woods. *Digital image processing*. Prentice Hall, 2008. 1

[13] D. K. Hammond and E. P. Simoncelli. Image denoising with an orientation-adaptive gaussian scale mixture model. In *IEEE ICIP*, 2006. 2

[14] N. Joshi, C. L. Zitnick, R. Szeliski, and D. Kriegman. Image deblurring and denoising using color priors. In *IEEE CVPR*, 2009. 2, 7

[15] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM Journal of Optimization*, 1998.

[16] A. B. Lee, D. Mumford, and J. Huang. Occlusion models for natural images: a statistical study of a scale-invariant dead leaves model. *IJCV*, 41:35–59, 2001. 2, 3

[17] A. Levin, R. Fergus, F. Durand, and W. T. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM TOG (SIGGRAPH)*, 2007. 1, 5

[18] C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang. Noise estimation from a single image. In *IEEE CVPR*, 2006. 3

[19] G. Matheron. *Random Sets and Integral Geometry*. John Wiley and Sons, 1975. 2, 3

[20] S. Roth and M. J. Black. Fields of experts: a framework for learning image priors. In *IEEE CVPR*, June 2005. 1

[21] S. Roth and M. J. Black. Steerable random fields. In *IEEE ICCV*, 2007. 2, 5, 7

[22] B. Russell, A. Torralba, K. Murphy, and W. T. Freeman. LabelMe: a database and web-based tool for image annotation. *IJCV*, 2008. 5

[23] S. S. Saquib, C. A. Bouman, and K. Sauer. ML parameter estimation for markov random fields with applications to bayesian tomography. *IEEE TIP*, 7(7):1029, 1998. 1, 2

[24] E. P. Simoncelli and E. H. Adelson. Noise removal via bayesian wavelet coring. In *IEEE ICIP*, 1996. 3

[25] C. V. Stewart. Robust parameter estimation in computer vision. *SIAM Reviews*, 41(3):513 – 537, Sept. 1999. 5

[26] M. F. Tappen, C. Liu, E. H. Adelson, and W. T. Freeman. Learning Gaussian conditional random fields for low-level vision. In *IEEE CVPR*, 2007. 2, 4, 7, 8

[27] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *IEEE ICCV*, 1998. 2

[28] M. Wainwright and E. P. Simoncelli. Scale mixtures of Gaussians and the statistics of natural images. In *NIPS*, 2000. 1

[29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE TIP*, 2004. 5

[30] Y. Weiss and W. T. Freeman. What makes a good model of natural images? In *IEEE CVPR*, 2007. 1

[31] M. Welling, G. Hinton, and S. Osindero. Learning sparse topographic representations with products of student-t distributions. In *NIPS*, 2002. 1

[32] S. C. Zhu, Y. Wu, and D. Mumford. Filters, random fields and maximum entropy (FRAME): Towards a unified theory for texture modeling. *IJCV*, 1998. 1